WHAT IS CLAIMED IS:

1.   A method of storing parameters of a deleted interpolation language model, the method comprising:

    obtaining a set of parameters for the deleted interpolation language model; and

    storing at least one parameter for the deleted interpolation language model as a parameter for a backoff language model.

2.   The method of claim 1 wherein storing at least one parameter for a deleted interpolation language model comprises storing an interpolated probability of a sequence of words as a probability for a sequence of words in the backoff language model.

3.   The method of claim 2 wherein storing an interpolated probability comprises determining that a relative frequency for the sequence of words is greater than a threshold before storing the interpolated probability as a probability.

4.   The method of claim 3 wherein the relative frequency is determined based on frequency counts that have fractional values.

5.   The method of claim 2 wherein storing an interpolated probability comprises determining that

the sequence of words forms a context for an n-gram in the backoff language model before storing the interpolated probability as a probability.

6.   The method of claim 1 wherein storing at least one parameter for a deleted interpolation language model comprises storing an interpolation weight for the deleted interpolation model as a backoff weight for the backoff language model.

7.   The method of claim 6 wherein storing the interpolation weight further comprises storing a sequence of words associated with the interpolation weight in a same entry as the interpolation weight.

8.   The method of claim 1 wherein obtaining a set of parameters comprises training a set of interpolation weights.

9.   The method of claim 8 wherein training a set of interpolation weights comprises training a separate weight for each of a set of frequency count ranges.

10.   The method of claim 1 wherein storing at least one parameter for the deleted interpolation language model comprises storing the at least one parameter to produce a data structure that conforms to the ARPA format for backoff language models.

11.   A computer-readable medium having computer-executable instructions for performing steps comprising:

      identifying a parameter for a deleted interpolation language model; and

      placing the parameter in a data structure as a backoff parameter for a backoff language model.

12.   The computer-readable medium of claim 11 wherein placing the parameter in a data structure comprises determining that the parameter should be included as part of the backoff language model.

13.   The computer-readable medium of claim 12 wherein determining that the parameter should be included as part of the backoff language model comprises determining that a frequency of a sequence of words in a training text exceeds a threshold.

14.   The computer-readable medium of claim 12 wherein determining that the parameter should be included as part of the backoff language model comprises determining that a sequence of words associated with the parameter forms a context in an n-gram stored in the data structure.

15.   The computer-readable medium of claim 11 wherein placing a parameter in a data structure

comprises placing an interpolated probability as a probability for an n-gram.

16.   The computer-readable medium of claim 11 wherein placing a parameter in a data structure comprises placing an interpolation weight as a backoff weight for a context.

17.   The computer-readable medium of claim 11 wherein the data structure conforms to the ARPA standard for backoff language models.

18.   A method for constructing a language model, the method comprising:

    using deleted interpolation to train parameters for a language model;

    storing at least some of the trained parameters in a data structure conforming to the ARPA format for backoff language models.

19.   The method of claim 18 wherein storing at least some of the trained parameters comprises storing parameters that are associated with sequences of words that appear in a training text more than a threshold amount.

20.   The method of claim 18 wherein storing at least some of the trained parameters comprises storing parameters that are associated with sequences

of words that appear as context words in an n-gram stored in the data structure.